
CURSO

ESTADÍSTICA

APLICADA CON R

(8ª Edición)

Organizado por:

Facultad de Ciencias de la

Universidad Autónoma de Madrid

Madrid, de septiembre a noviembre de 2026

CURSO DE FORMACIÓN CONTINUA: ESTADÍSTICA APLICADA CON R

PRESENTACIÓN

El **Departamento de Matemáticas** de la Facultad de Ciencias de la Universidad Autónoma de Madrid organiza, conjuntamente con el **Departamento de Bioestadística de GEICAM** e investigadores de **FUNDAE**, un programa de formación continua en **Estadística Aplicada con el software R**.

El objetivo del curso es ofrecer una **formación práctica y rigurosa para analizar datos reales con R**, desde las primeras fases de importación, limpieza, transformación y visualización, hasta la aplicación de modelos estadísticos, técnicas de *machine learning*, métodos especializados e informes reproducibles. El curso no se plantea únicamente como una sucesión de técnicas estadísticas, sino como un itinerario orientado a resolver problemas de análisis de datos en **contextos académicos, científicos, biomédicos y profesionales**.

El programa está compuesto por **diferentes módulos** que pueden ser cursados en su totalidad o de manera independiente. Esta estructura modular permite construir distintos recorridos formativos: desde una introducción práctica al uso de R hasta módulos más avanzados sobre visualización interactiva, análisis multivariante, bioestadística, estadística bayesiana, machine learning, inteligencia artificial aplicada a la docencia y la investigación, y análisis de datos funcionales.

Los contenidos abarcan desde herramientas básicas de análisis descriptivo e inferencia estadística hasta **métodos avanzados y especializados**, como el análisis multivariante, datos complejos y de alta dimensión, estadística bayesiana, técnicas de *machine learning* y el uso de herramientas de inteligencia artificial aplicadas a la docencia y la investigación. El programa se estructura en **nueve módulos**, diseñados para resultar de interés a distintos perfiles profesionales. Aunque algunos contenidos aparecen de forma transversal en varios módulos, estos se abordan desde perspectivas prácticas y complementarias, lo que permite al alumnado construir itinerarios formativos adaptados a sus necesidades e intereses.

R es actualmente el software **de referencia** en numerosos centros de investigación y universidades, y su uso se ha extendido de manera significativa en el ámbito empresarial. Se trata de un **software libre y gratuito** que incorpora con rapidez los **métodos más avanzados de análisis de datos**. Además, su **lenguaje de programación**, sencillo y flexible permite un control completo y reproducible de los análisis estadísticos realizados.

El módulo 1 introduce el lenguaje R y las herramientas básicas para el análisis estadístico, por lo que está especialmente recomendado para quienes no tengan experiencia previa con este entorno. El resto de los módulos pueden ser cursados una vez que se ha asistido al de introducción, o también por alumnos que tengan ya conocimientos de R y de estadística básica.

DIRIGIDO A

Este programa de formación continua va dirigido a **profesionales** de distintos ámbitos, **investigadores, profesores y alumnos**, tanto de la UAM como de otras instituciones, que quieran conocer el software R y su aplicación práctica con diferentes técnicas estadísticas.

El curso está especialmente orientado a personas que necesiten trabajar con datos reales, interpretar resultados estadísticos y comunicar conclusiones de forma clara, rigurosa y reproducible. Los módulos pueden cursarse de forma independiente o combinarse según los intereses y necesidades de cada participante.

Como orientación, algunos itinerarios posibles:

- **Itinerario de iniciación al análisis de datos con R**, dirigido a quienes quieran adquirir una base sólida en el uso de R, la manipulación de datos, la visualización y la elaboración de informes reproducibles. Incluye módulos 1, 2 y 3.
- **Itinerario para investigación biomédica y clínica**, orientado a investigadores, profesionales sanitarios y perfiles técnicos que necesiten aplicar R al análisis de datos biomédicos, clínicos o epidemiológicos. Incluye módulos 1, 5, 6 y 7.
- **Itinerario de ciencia de datos y aprendizaje automático**, pensado para participantes interesados en modelos predictivos, clasificación, machine learning con tidymodels, automatización de flujos de análisis y uso crítico de herramientas de inteligencia artificial. Incluye módulos 3, 4, 5, 8 y 9.
- **Itinerario de automatización, comunicación e IA**, dirigido a participantes interesados en automatizar tareas, crear herramientas interactivas, mejorar la comunicación de resultados y explorar el uso crítico de la inteligencia artificial en contextos docentes, investigadores o profesionales. Incluye módulos 2, 3, 4 y 9.
- **Itinerario completo en estadística aplicada con R**, dirigido a quienes quieran aprovechar el programa como una experiencia formativa global, práctica y progresiva. Este recorrido permite desarrollar una forma de trabajo sólida con datos reales: preparar y explorar información, detectar patrones, comparar grupos, estudiar asociaciones, construir modelos, evaluar predicciones e interpretar la incertidumbre de los resultados. Es la opción más adecuada para quienes buscan una visión amplia y conectada del análisis de datos con R, combinando rigor estadístico, herramientas modernas y aplicaciones en contextos científicos, docentes y profesionales.

ESTRUCTURA

Módulos	Fechas	Núm. Horas	Precio (*)	Créditos
1. Introducción a R para análisis de datos	21, 22 y 23 Septiembre 2026	15	180€ (162€)	1
2. Visualización interactiva de datos con el paquete Shiny	29 y 30 Septiembre 2026	10	120€ (108€)	1
3. Programar mejor con R: automatización, APIs y análisis reproducibles	7 y 8 Octubre 2026	10	120€ (108€)	1
4. Agentes de IA en R para docencia, investigación y revisión bibliográfica	14 y 15 Octubre 2026	10	120€ (108€)	1
5. Modelos estadísticos para ciencia de datos con R: regresión, clasificación y reducción de la dimensión	20, 21 y 22 Octubre 2026	15	180€ (162€)	1
6. Estadística aplicada a la investigación biomédica con R	27, 28 y 29 Octubre 2026	20	240€ (216€)	2
7. Estadística bayesiana aplicada con R: modelos, incertidumbre y decisiones	3, 4 y 5 Noviembre 2026	15	180€ (162€)	1
8. Análisis de curvas, señales y datos funcionales con R	10, 11 y 12 Noviembre 2026	15	180€ (162€)	1
9. Machine Learning con R y tidymodels	16, 17, 18, 19 y 20 Noviembre 2026	30	360€ (324€)	3

El coste máximo del curso es de 1680€

(*) Descuento 10% para estudiantes con matrícula en vigor, AlumniUAM+, AlumniUAM+Plus y Amigos de la UAM

CRÉDITOS

Se podrán reconocer hasta 3 créditos.

DIRECTORES

- José Luis Torrecilla Nogueras (Profesor Permanente en el Dpto. de Matemáticas de la UAM)
- Hristo Inouzhe Valdes (Profesor AD en el Dpto. de Matemáticas de la UAM)
- Jesús Herranz Valera (Bioestadístico en GEICAM)

PERSONAL DOCENTE

- Carlos Arias Alcaide (Científico de datos en FUNDAE)
- José Ramón Berrendero Díaz (Catedrático en el Dpto. de Matemáticas de la UAM)
- Jesús Herranz Valera (Bioestadístico en GEICAM)
- Hristo Inouzhe Valdes (Profesor Ayudante Doctor en el Dpto. de Matemáticas de la UAM)
- Martín Sánchez Signorini (Becario FPU-MICIU)
- José Luis Torrecilla Nogueras (Profesor Permanente en el Dpto. de Matemáticas de la UAM)

INSCRIPCIÓN Y MATRICULACIÓN

Las **inscripciones** en los diferentes módulos se realizarán en la **página web de la Fundación de la Universidad Autónoma de Madrid**:

<http://matriculas.fuam.es/matriculauam/Convocatorias.action>

Para cada módulo, los **plazos de matriculación** comenzarán el día 22 de junio de 2026 y se extenderán hasta tres días antes al comienzo del módulo.

Una vez realizada la matriculación a través de la página web de la UAM, se deberá realizar **el pago** de la misma **antes del inicio del módulo**, en un **plazo inferior a 15 días**. Si en ese plazo no se ha ingresado dicho importe, se procederá a la **anulación de la matrícula**.

Las **plazas** en cada módulo se adjudicarán por **orden de ingreso** del importe de la matrícula.

Los alumnos interesados en el curso completo, tendrán que realizar la matrícula en cada uno de los módulos por separado.

LUGAR DE CELEBRACIÓN

Todos los **módulos** se impartirán en la **Facultad de Ciencias de la UAM**.

Es **obligatorio** que los participantes traigan su propio **ordenador portátil**.

Las **aulas** donde se impartirán las clases se comunicarán a los alumnos días antes de comenzar cada módulo.

BECAS

Entre los alumnos matriculados en cada módulo se concederán hasta **dos becas** (en función del número final de inscritos) que cubrirán el total del módulo correspondiente.

Los **criterios de selección** de becarios son:

1. Criterios económicos de concesión de becas del Ministerio de Educación y situación de desempleo: (60%)
2. Expediente académico y/o experiencia profesional relacionados con el contenido del título o que contribuyan a su mayor aprovechamiento (20%)
3. Motivación para cursar los estudios (20%)

Las **solicitudes de las becas** se harán por correo electrónico dirigido al personal docente del módulo. En el correo se proporcionará información sobre los puntos anteriores, indicando la situación laboral del alumno: desempleado, personal fijo, contratado eventual o becario. En caso necesario, se podrá solicitar una declaración jurada.

La **asignación de las becas** se realizará después de finalizada la matrícula de cada módulo, y el importe de la matrícula será devuelto posteriormente a los alumnos que hayan sido becados.

El profesorado del curso decidirá en cada módulo si en lugar de conceder becas completas, se conceden el doble de **medias becas**, con la devolución de la **mitad del importe** de la matrícula a cada uno de los alumnos becados.

El profesorado del curso dará **prioridad** en cada módulo a los **alumnos que no hayan sido becados** en módulos anteriores.

INFORMACIÓN ADICIONAL

Este curso forma parte de los Cursos de Formación Continua organizados por la Fundación de la Universidad Autónoma de Madrid, FUAM, y se puede consultar en: <https://fuam.es/cursos-cortos/>

MÓDULO 1: INTRODUCCIÓN A R PARA ANÁLISIS DE DATOS

PRESENTACIÓN Y OBJETIVOS

Este módulo ofrece una **introducción práctica al uso de R** para el análisis de datos, dirigida a personas que no han trabajado previamente con este entorno o que necesitan consolidar una base para utilizarlo con autonomía. El curso está especialmente orientado a estudiantes de posgrado, investigadores y profesionales que quieren comenzar a aplicar R en tareas habituales de **análisis, visualización e interpretación de datos**.

A lo largo del módulo se presentarán los elementos básicos del lenguaje R y del entorno **RStudio**, incluyendo la creación y ejecución de **scripts**, los principales **tipos de objetos**, la importación y manipulación de datos, la elaboración de gráficos y la realización de análisis estadísticos sencillos. Además, se introducirán herramientas del ecosistema **tidyverse**, especialmente **dplyr** y **ggplot2**, por su utilidad práctica en análisis de datos.

El módulo también incorporará una primera aproximación a la elaboración de **informes reproducibles con R Markdown**, que permite integrar texto, código, resultados, tablas y gráficos en un mismo documento. De este modo, los participantes aprenderán a organizar, documentar y comunicar sus resultados de forma clara y reproducible.

El enfoque será progresivo y aplicado. Las sesiones combinarán explicaciones conceptuales, ejemplos guiados y ejercicios prácticos, con el objetivo de que los participantes adquieran una base sólida para continuar con el resto de módulos del curso o para utilizar R en sus propios proyectos académicos, científicos o profesionales.

INSCRIPCIÓN Y DATOS PRÁCTICOS

MÓDULO	Introducción a R para análisis de datos
Docentes	Martín Sánchez y José Luis Torrecilla (Dpto. Matemáticas, UAM) jose Luis.torrecilla@uam.es
Fechas del Curso	21, 22 y 23 de Septiembre de 2026. Horario: 15:00 – 20:00
Inscripciones	Hasta el 17/Sept/2026 en http://matriculas.fuam.es/matriculauam/Convocatorias.action
Lugar de celebración	Universidad Autónoma de Madrid Ctra. Madrid-Colmenar Viejo (M-607), km 15. Madrid. Facultad de Ciencias (C/ Francisco Tomás y Valiente, 7) (por confirmar) http://www.uam.es
Nº Plazas	16 plazas. Mínimo de 8 asistentes para impartirse el curso
Precio	180€ (162€ estudiantes con matrícula en vigor, AlumniUAM+, AlumniUAM+Plus y Amigos de la UAM)
Créditos	1 ECTS

Requisitos previos	No se requieren conocimientos previos de R. Para el máximo aprovechamiento del módulo, es conveniente tener conocimientos elementales de estadística (en caso de duda contacten con los profesores del módulo)
Observaciones	Los contenidos teóricos, conjuntos de datos y código necesarios para el seguimiento del módulo se harán accesibles en documentos en red. Los participantes deben traer su propio ordenador portátil . Se enviarán indicaciones sobre el software a instalar antes del inicio del módulo.

DESCRIPCIÓN DETALLADA

El módulo está dividido en 3 sesiones de 5 horas.

El curso está concebido como una introducción progresiva y práctica al uso del software R para el análisis estadístico de datos. A lo largo del módulo, los participantes se familiarizarán con el entorno de trabajo *RStudio*, así como con los principales tipos de objetos y estructuras de datos, sentando las bases necesarias para abordar análisis estadísticos de forma autónoma.

Los contenidos combinan la exposición de conceptos estadísticos fundamentales con el aprendizaje del manejo del lenguaje R, poniendo un énfasis especial en la visualización y exploración de datos, y la elaboración de análisis reproducibles. Las sesiones se apoyan en ejemplos prácticos y en la realización de ejercicios guiados, lo que permite afianzar los conocimientos adquiridos y adquirir soltura en el uso del software para el análisis de datos.

1. Primeros pasos en R: consola, scripts, interfaz RStudio
2. Objetos y estructuras básicas de datos en R
3. Importación de conjuntos de datos
4. Introducción a la descripción y visualización de datos
5. Ejemplos sencillos de simulación e inferencia estadística
6. Introducción a la programación en R: funciones y estructuras básicas
7. Manipulación de *data frames* con *dplyr*
8. Representación gráfica con *ggplot2*
9. Elaboración de informes dinámicos con *R Markdown*

MÓDULO 2: VISUALIZACIÓN INTERACTIVA DE DATOS CON EL PAQUETE SHINY

PRESENTACIÓN Y OBJETIVOS

Durante la realización de análisis estadísticos, la cantidad de información que generamos y tenemos que gestionar puede ser enorme. A la hora de presentar dicha información, si no se usa el método adecuado, la tarea se vuelve tediosa. Este problema ha dado lugar a un nuevo campo de trabajo llamado **Visualización de Datos** (Data Visualization).

El paquete **Shiny** nos adentra a los usuarios de R en la Visualización de Datos, facilitándonos la creación de **aplicaciones interactivas**. En estas aplicaciones podemos mostrar de manera ordenada gráficos, resúmenes de información, datos, etc., disponiendo de menús, casillas para marcar, desplegables, deslizadores de valores, entre otras opciones.

El objetivo principal de este módulo es que los asistentes aprendan a **reproducir, crear y configurar sus propias aplicaciones interactivas con el paquete Shiny** de R, desarrollando un pensamiento razonado y creativo.

INSCRIPCIÓN Y DATOS PRÁCTICOS

MÓDULO	Visualización Interactiva de Datos con el paquete Shiny
Docentes	Carlos Arias Alcaide (Científico de datos en FUNDAE) carlos.arias.cienciadedatos@gmail.com
Fechas del Curso	29 y 30 de Septiembre de 2026. Horario: 16:00 – 21:00
Inscripciones	Hasta el 25/Sept/2026 en http://matriculas.fuam.es/matriculauam/Convocatorias.action
Lugar de celebración	Universidad Autónoma de Madrid Ctra. Madrid-Colmenar Viejo (M-607), km 15. Madrid. Facultad de Ciencias (C/ Francisco Tomás y Valiente, 7) (por confirmar) http://www.uam.es
Nº Plazas	16 plazas. Mínimo de 8 asistentes para impartirse el curso
Precio	120€ (108€ estudiantes con matrícula en vigor, AlumniUAM+, AlumniUAM+Plus y Amigos de la UAM)
Créditos	1 ECTS
Requisitos previos	Son necesarias nociones básicas de R y RStudio. (en caso de duda contacten con los profesores del módulo)
Observaciones	Los asistentes contarán con unos apuntes para que puedan seguir las explicaciones, repasar los contenidos y consultar posteriormente tras finalizar el curso. Es obligatorio que los participantes traigan su propio ordenador portátil con el software R y RStudio instalados previamente.

DESCRIPCIÓN DETALLADA

El módulo está dividido en 2 sesiones de 5 horas cada una, en las que se alternan explicaciones teóricas con la aplicación en un ejemplo y un ejercicio práctico guiado.

En la sesión 1 se explicarán los objetivos del curso, la terminología empleada y la estructura de la aplicación, haciendo hincapié en la parte visible (*front-end* o interfaz de usuario). En la sesión 2, se continuará con la estructura de la aplicación, centrándose en la parte no visible (*back-end* o servidor), se explicarán diferentes formas de compartir nuestras aplicaciones con otros usuarios y se terminará desarrollando algunos casos prácticos.

1. Introducción y objetivos del curso
2. Términos
3. Estructura de la aplicación: *front-end* o interfaz de usuario
4. Estructura de la aplicación: *back-end* o servidor
5. Compartir la aplicación con otros usuarios
6. Estudio, desarrollo y resolución de casos prácticos

MÓDULO 3: PROGRAMAR MEJOR CON R: AUTOMATIZACIÓN, APIS Y ANÁLISIS REPRODUCIBLES

PRESENTACIÓN Y OBJETIVOS

Este módulo está dirigido a personas que ya han tenido una primera aproximación a R y quieren aprender a **escribir código más claro, ordenado, eficiente y reutilizable**. Asimismo, se introducirán algunas prácticas y herramientas más avanzadas. No se presupone experiencia previa en programación avanzada. El curso parte de conceptos básicos y los desarrolla de forma progresiva, con ejemplos orientados al análisis de datos y a situaciones habituales en investigación.

El objetivo principal es ayudar a los participantes a **pasar de scripts aislados a análisis mejor organizados y reproducibles**. Para ello se introducirán ideas básicas como la modularidad, el diseño de funciones, la reutilización de código y la separación entre datos, código y resultados. También se trabajará el flujo del programa con condiciones, bucles y herramientas para repetir operaciones de forma ordenada y **automatizar** tareas frecuentes, con especial atención al paquete *purrr* dentro del ecosistema *tidyverse*.

A lo largo del módulo se abordará la lectura de múltiples archivos, la organización de proyectos, la detección de errores y algunas buenas prácticas de **rendimiento**, incluyendo vectorización y paralelización básica. El curso incorporará además una introducción práctica a herramientas actuales como **APIs**, formato **JSON**, **Git** para control de versiones y uso crítico de herramientas de **inteligencia artificial** como apoyo a la programación.

El enfoque será eminentemente práctico, mediante explicaciones conceptuales, ejemplos guiados y ejercicios en R, con el objetivo de que los participantes puedan entender, modificar y mejorar sus propios scripts en proyectos académicos, científicos o profesionales.

INSCRIPCIÓN Y DATOS PRÁCTICOS

MÓDULO	Programar mejor con R: automatización, APIs y análisis reproducibles
Docentes	José Luis Torrecilla y Hristo Inouzhe Valdes (Dpto. Matemáticas, UAM) joseluis.torrecilla@uam.es
Fechas del Curso	7 y 8 de Octubre de 2026. Horario: 15:00 – 20:00
Inscripciones	Hasta el 3/Oct/2026 en http://matriculas.fuam.es/matriculauam/Convocatorias.action
Lugar de celebración	Universidad Autónoma de Madrid Ctra. Madrid-Colmenar Viejo (M-607), km 15. Madrid. Facultad de Ciencias (C/ Francisco Tomás y Valiente, 7) (por confirmar) http://www.uam.es
Nº Plazas	16 plazas. Mínimo de 8 asistentes para impartirse el curso
Precio	120€ (108€ estudiantes con matrícula en vigor, AlumniUAM+, AlumniUAM+Plus)

	y Amigos de la UAM)
Créditos	1 ECTS
Requisitos previos	Se requieren nociones básicas de R y RStudio, equivalentes a las adquiridas en un curso introductorio: creación y ejecución de scripts, manejo básico de objetos y data frames, e importación sencilla de datos. No se requieren conocimientos previos de programación avanzada, APIs, Git ni herramientas de inteligencia artificial. En caso de duda, se recomienda contactar con el profesorado del módulo.
Observaciones	Los contenidos teóricos, conjuntos de datos y código necesarios para el seguimiento del módulo se harán accesibles en documentos en red. Los participantes deben traer su propio ordenador portátil . Se enviarán indicaciones sobre el software a instalar antes del inicio del módulo.

DESCRIPCIÓN DETALLADA

El módulo está dividido en 2 sesiones de 5 horas.

El curso introduce de forma progresiva conceptos y herramientas para escribir código en R más claro, reutilizable y reproducible. Se trabajarán ideas básicas de programación, como la modularidad, el diseño de funciones, las estructuras de control y la automatización de tareas repetitivas.

Las sesiones combinarán explicaciones conceptuales, ejemplos desarrollados en R y ejercicios guiados. Se utilizarán herramientas de R base y del ecosistema *tidyverse*. También se abordarán buenas prácticas para organizar proyectos, trabajar con múltiples archivos, detectar errores y mejorar la eficiencia de algunos análisis.

Además, se introducirán de forma práctica herramientas actuales como APIs, datos en formato JSON, Git, vectorización, paralelización básica y uso crítico de IA como apoyo a la programación.

1. Conceptos básicos de programación: modularidad, funciones y reutilización de código
2. Paradigmas de programación en R: ideas generales
3. De scripts lineales a análisis reproducibles
4. Control de flujo: condiciones, bucles y aplicación repetida de operaciones
5. Automatización de tareas frecuentes: *purrr*
6. Errores frecuentes, depuración básica y revisión del código
7. Introducción al uso de APIs y datos en formato JSON
8. Uso crítico de herramientas de IA como apoyo a la programación
9. Introducción a Git y control de versiones
10. Rendimiento en R: vectorización y paralelización básica

MÓDULO 4: AGENTES DE IA EN R PARA DOCENCIA, INVESTIGACIÓN Y REVISIÓN BIBLIOGRÁFICA

PRESENTACIÓN Y OBJETIVOS

Los **modelos de lenguaje** (LLMs) se están convirtiendo rápidamente en una herramienta práctica para investigar y enseñar, pero el salto de verdad llega cuando dejas de “chatear” y empiezas a **construir agentes**: sistemas capaces de buscar, leer, razonar con tus propios materiales y generar resultados fiables que puedas reutilizar. En este curso aprenderás qué son los LLMs y los agentes (de forma clara y sin exageración), qué pueden y qué no pueden hacer, y cómo usarlos de manera útil y confiable en el trabajo académico. Trabajaremos en R y RStudio, con requisitos previos mínimos, y nos centraremos en **flujos reproducibles** en lugar de apps tipo caja negra.

Al terminar, habrás construido dos agentes funcionales que podrás adaptar a tus necesidades: un agente Profesor/Tutor anclado en tus apuntes (que responde citando tus materiales y sabe decir “no aparece” cuando corresponde) y un agente Explorador de investigación que busca en arXiv, filtra artículos y produce mini-reseñas estructuradas y recomendaciones. Por el camino aprenderás prácticas modernas para mejorar la fiabilidad—como la generación aumentada por recuperación (RAG), salidas estructuradas y comprobaciones de evaluación sencillas—para que tus agentes se conviertan en herramientas robustas **para preparar clases, supervisar proyectos y acelerar revisiones bibliográficas**.

INSCRIPCIÓN Y DATOS PRÁCTICOS

MÓDULO	Agentes de IA en R para docencia, investigación y revisión bibliográfica
Docentes	Hristo Inouze Valdes (Dpto. Matemáticas, UAM) hristo.inouze@uam.es
Fechas del Curso	14 y 15 de Octubre de 2026. Horario: 15:00 – 20:00
Inscripciones	Hasta el 10/Oct/2026 en http://matriculas.fuam.es/matriculauam/Convocatorias.action
Lugar de celebración	Universidad Autónoma de Madrid Ctra. Madrid-Colmenar Viejo (M-607), km 15. Madrid. Facultad de Ciencias (C/ Francisco Tomás y Valiente, 7) (por confirmar) http://www.uam.es
Nº Plazas	16 plazas. Mínimo de 8 asistentes para impartirse el curso
Precio	120€ (108€ estudiantes con matrícula en vigor, AlumniUAM+, AlumniUAM+Plus y Amigos de la UAM)
Créditos	1 ECTS
Requisitos previos	Son necesarias nociones básicas de R y RStudio. (en caso de duda contacten con los profesores del módulo)
Observaciones	Es obligatorio que los participantes traigan su propio ordenador portátil con el software R y RStudio instalados previamente.

DESCRIPCIÓN DETALLADA

El módulo está dividido en 2 sesiones de 5 horas.

En cada sesión se explican los principales conceptos, tanto estadísticos como relativos al funcionamiento del programa, y se ilustran de manera práctica con ejemplos realizados con R. Finalmente, se proponen ejercicios a los alumnos para que puedan practicar los conocimientos adquiridos.

1. Qué son los LLMs (y qué no son)
2. De chatbot a agente: herramientas, memoria y bucles de control sencillos
3. Patrones de prompting que funcionan en flujos de trabajo con R
4. Salidas estructuradas: esquemas, parseo y validación en R
5. RAG con tus materiales: troceado, embeddings, recuperación y citas
6. Construcción de un agente Docente/Tutor a partir de apuntes (respuestas fundamentadas + rechazo)
7. Construcción de un agente “scout” para arXiv: buscar, filtrar, ordenar, resumir, recomendar
8. Kit de fiabilidad: control de alucinaciones, guardarraíles y evaluaciones rápidas
9. Pipelines reproducibles: scripts, plantillas y diseño “agente como función”
10. Uso responsable en educación: privacidad, atribución e integridad académica

MÓDULO 5: MODELOS ESTADÍSTICOS PARA CIENCIA DE DATOS CON R: REGRESIÓN, CLASIFICACIÓN Y REDUCCIÓN DE DIMENSIÓN

PRESENTACIÓN Y OBJETIVOS

Este módulo tiene como objetivo proporcionar una formación práctica y aplicada en algunos de los **modelos estadísticos fundamentales** que se utilizan actualmente en ciencia de datos para analizar información con múltiples variables. El curso está orientado a **comprender, aplicar e interpretar** técnicas que permiten resumir datos complejos, estudiar relaciones entre variables, detectar patrones, formar grupos y construir reglas de clasificación mediante el software R.

A lo largo del módulo se introducirán herramientas básicas para el análisis de datos multivariantes, prestando especial atención a su utilidad en **problemas reales de investigación**. Se abordarán técnicas de **reducción de dimensión**, útiles para representar datos con muchas variables y extraer estructuras relevantes; **modelos de regresión**, que permiten estudiar y explicar la relación entre una variable de interés y un conjunto de variables explicativas; y métodos de **clasificación y clustering**, orientados a identificar perfiles, patrones o grupos de observaciones.

El enfoque del curso será eminentemente aplicado. Las técnicas se presentarán desde una perspectiva estadística, pero siempre acompañadas de ejemplos desarrollados en R, interpretación de resultados y discusión de sus usos y limitaciones. Se utilizarán paquetes habituales de R para análisis estadístico y modelización, incluyendo, cuando proceda, herramientas como *caret* para **organizar flujos básicos** de entrenamiento, validación y comparación de modelos. El objetivo final es que los participantes sean capaces de seleccionar y aplicar herramientas adecuadas para sus propios datos, especialmente en contextos de investigación científica, experimental o profesional en los que es habitual trabajar con muchas variables y necesitar **resultados interpretables, reproducibles y comunicables**.

INSCRIPCIÓN Y DATOS PRÁCTICOS

MÓDULO	Modelos estadísticos para ciencia de datos con R: regresión, clasificación y reducción de la dimensión
Docentes	José Ramón Berrendero y José Luis Torrecilla (Dpto. Matemáticas, UAM) jose Luis.torrecilla@uam.es
Fechas del Curso	20, 21 y 22 de Octubre de 2026. Horario: 15:00 – 20:00
Inscripciones	Hasta el 16/Oct/2026 en http://matriculas.fuam.es/matriculauam/Convocatorias.action
Lugar de celebración	Universidad Autónoma de Madrid Ctra. Madrid-Colmenar Viejo (M-607), km 15. Madrid. Facultad de Ciencias (C/ Francisco Tomás y Valiente, 7) (por confirmar) http://www.uam.es
Nº Plazas	16 plazas. Mínimo de 8 asistentes para impartirse el curso
Precio	180€ (162€ estudiantes con matrícula en vigor, AlumniUAM+, AlumniUAM+Plus y Amigos de la UAM)
Créditos	1 ECTS
Requisitos previos	Para el máximo aprovechamiento del módulo se requieren conocimientos

	básicos de R y de inferencia estadística (en caso de duda contacten con los profesores del módulo)
Observaciones	Los contenidos teóricos, conjuntos de datos, guiones de prácticas y código necesario para el seguimiento del módulo se pondrán a disposición de los participantes a través de una plataforma o repositorio en línea. Los participantes deberán traer su propio ordenador portátil. Antes del inicio del módulo se enviarán indicaciones detalladas sobre el software y los paquetes de R necesarios para realizar las prácticas. El módulo tendrá un enfoque aplicado y reproducible. Se trabajará con ejemplos y ejercicios guiados orientados al análisis de datos con múltiples variables, prestando atención tanto a la implementación en R como a la interpretación de los resultados.

DESCRIPCIÓN DETALLADA

El módulo está dividido en 3 sesiones de 5 horas.

En cada sesión se explican los principales conceptos, tanto estadísticos como relativos al funcionamiento del programa. Las clases se apoyan en ejemplos prácticos y en la realización de ejercicios guiados, lo que permite afianzar los conocimientos adquiridos y adquirir soltura en el uso del software para el análisis de datos.

1. Descripción de datos y distribuciones multivariantes
2. Reducción de la dimensión y representación de datos complejos
3. *Clustering* o análisis de conglomerados: detección de grupos y perfiles en los datos
4. Clasificación supervisada: introducción al análisis discriminante
5. Modelos de regresión para explicación y predicción
6. Selección de variables e interpretación de modelos
7. Diagnóstico del modelo y tratamiento de datos atípicos
8. Introducción a flujos básicos de modelización y validación con R: *caret*

MÓDULO 6: ESTADÍSTICA APLICADA A LA INVESTIGACIÓN BIOMÉDICA CON R

PRESENTACIÓN Y OBJETIVOS

El objetivo principal de este módulo es proporcionar a los alumnos los conocimientos de las técnicas estadísticas más utilizadas dentro del campo de las ciencias de la salud, de la epidemiología y de la investigación biomédica en general.

El módulo incluye las técnicas estadísticas más sencillas para analizar **datos clínicos y epidemiológicos**, como son el análisis descriptivo de datos y la inferencia básica, pero la parte más importante está dedicada a los modelos de regresión. Se ha dado especial importancia a las técnicas más usadas en la investigación biomédica, que son la **regresión logística** y la **regresión de Cox**, para analizar **datos de supervivencia**. Los modelos de regresión se explican desde un punto de vista práctico, y se han incluido temas avanzados como: el tratamiento de las **variables de confusión**, el análisis e interpretación de **interacciones**, las estrategias para la **construcción de modelos** de regresión multivariantes y el análisis de la **capacidad predictiva** de los modelos.

INSCRIPCIÓN Y DATOS PRÁCTICOS

MÓDULO	Estadística Aplicada a la Investigación Biomédica con R
Docentes	Jesús Herranz Valera (Bioestadístico de GEICAM) jherranzvalera@gmail.com
Fechas del Curso	27, 28 y 29 de Octubre de 2026. Horario: 9:00 – 18:00
Inscripciones	Hasta el 23/Oct/2026 en http://matriculas.fuam.es/matriculauam/Convocatorias.action
Lugar de celebración	Universidad Autónoma de Madrid Ctra. Madrid-Colmenar Viejo (M-607), km 15. Madrid. Facultad de Ciencias (C/ Francisco Tomás y Valiente, 7) (por confirmar) http://www.uam.es
Nº Plazas	16 plazas. Mínimo de 8 asistentes para impartirse el curso
Precio	240€ (216€ estudiantes con matrícula en vigor, AlumniUAM+, AlumniUAM+Plus y Amigos de la UAM)
Créditos	2 ECTS
Requisitos previos	Para el máximo aprovechamiento del módulo se requieren conocimientos básicos de R, y conocimientos básicos de inferencia estadística (en caso de duda contacten con los profesores del módulo)
Observaciones	Se entregará un documento con todas las diapositivas del curso y todos los scripts y ficheros de datos Es obligatorio que los participantes traigan su propio ordenador portátil , y que se haya instalado el software R previamente

DESCRIPCIÓN DETALLADA

El módulo está dividido en 18 sesiones, con una duración aproximada de 1 hora cada una.

En cada sesión se explican los conceptos estadísticos teóricos ilustrados desde un punto de vista práctico con ejemplos realizados con R, explicando los paquetes y funcionalidades más importantes del software para cada técnica. Finalmente, se proponen ejercicios a los alumnos para que puedan practicar los conocimientos adquiridos.

1. Tratamiento de datos con R y *tidyverse*. Gráficos
2. Estadística descriptiva
3. Análisis de tablas de contingencia
4. Inferencia básica con variables continuas
5. Análisis de la varianza
6. Análisis de correlación
7. Regresión lineal
8. Regresión logística
9. Variables de confusión e interacciones
10. Construcción de un modelo de regresión logística
11. Modelos predictivos. Análisis de curvas ROC
12. Análisis de supervivencia
13. Regresión de Cox
14. Modelos predictivos. Análisis de curvas ROC
15. Análisis de Medidas Repetidas

MÓDULO 7: ESTADÍSTICA BAYESIANA APLICADA CON R: MODELOS, INCERTIDUMBRE Y DECISIONES

PRESENTACIÓN Y OBJETIVOS

Este curso presenta la estadística bayesiana como un **flujo de trabajo práctico** para modelar incertidumbre, analizar datos reales y comunicar resultados de forma clara. El tema central será entender un modelo bayesiano como un **mecanismo generativo**: primero pensamos qué tipo de datos podría producir el modelo, después incorporamos los datos observados y finalmente evaluamos si las predicciones del modelo son razonables. El asistente aprenderá a interpretar *priors*, verosimilitudes, distribuciones posteriores y distribuciones predictivas, prestando especial atención a la diferencia entre **incertidumbre sobre parámetros, incertidumbre predictiva y cantidades no observadas**.

El curso se centra en el uso aplicado de R para **construir, simular, ajustar y criticar modelos bayesianos**. Partiremos de ejemplos sencillos para introducir la probabilidad como incertidumbre y la regla de Bayes, y desarrollaremos una regresión lineal bayesiana completa como hilo conductor. A partir de ahí, ampliaremos el enfoque hacia modelos más realistas: modelos lineales generalizados para datos binarios y de conteo, modelos jerárquicos con *pooling* parcial, datos faltantes, supervivencia, inferencia causal y modelos flexibles como splines, procesos gaussianos y mezclas. El objetivo no es cubrir todos estos temas en profundidad teórica, sino proporcionar una **visión estructurada y aplicada** de cómo se integran dentro del flujo de trabajo bayesiano moderno.

INSCRIPCIÓN Y DATOS PRÁCTICOS

MÓDULO	Estadística bayesiana aplicada con R: modelos, incertidumbre y decisiones
Docentes	Hristo Inouze Valdes (Dpto. Matemáticas, UAM) hristo.inouze@uam.es
Fechas del Curso	3, 4 y 5 de Noviembre de 2026. Horario: 15:00 – 20:00
Inscripciones	Hasta el 30/Oct/2026 en http://matriculas.fuam.es/matriculauam/Convocatorias.action
Lugar de celebración	Universidad Autónoma de Madrid Ctra. Madrid-Colmenar Viejo (M-607), km 15. Madrid. Facultad de Ciencias (C/ Francisco Tomás y Valiente, 7) (por confirmar) http://www.uam.es
Nº Plazas	16 plazas. Mínimo de 8 asistentes para impartirse el curso
Precio	180€ (162€ estudiantes con matrícula en vigor, AlumniUAM+, AlumniUAM+Plus y Amigos de la UAM)
Créditos	1 ECTS
Requisitos previos	Para el máximo aprovechamiento del módulo, se requiere conocer al menos los rudimentos del lenguaje R y de la probabilidad (en caso de duda contacten con los profesores del módulo)
Observaciones	Es obligatorio que los participantes traigan su propio ordenador portátil , y que se haya instalado el software R previamente

DESCRIPCIÓN DETALLADA

El módulo está dividido en 3 sesiones de 5 horas.

En cada sesión se explican los principales conceptos, tanto estadísticos como relativos al uso de R, y se ilustran mediante ejemplos reproducibles en formato Quarto/revealjs. Finalmente, se proponen ejercicios y actividades para que los alumnos puedan practicar los conocimientos adquiridos sobre datos simulados y reales.

1. La probabilidad como medida de incertidumbre: incertidumbre epistémica, variabilidad aleatoria y cantidades no observadas
2. Flujo de trabajo bayesiano: modelo generativo, distribución a priori, verosimilitud, distribución posterior y predicción
3. Simulación predictiva previa: cómo comprobar si los *priors* y el modelo generan datos plausibles
4. Regresión lineal bayesiana como ejemplo completo: interpretación de parámetros, incertidumbre sobre la media e incertidumbre predictiva
5. Introducción computacional: Monte Carlo, MCMC, HMC, aproximación variacional e interpretación práctica de los diagnósticos
6. Evaluación del modelo: chequeos predictivos a posteriori, discrepancias relevantes, comparación predictiva, WAIC y LOO
7. Modelos lineales generalizados: regresión logística, modelos de conteo, tasas, *offsets* y sobredispersión
8. Modelos jerárquicos y *pooling* parcial para datos agrupados: centros, escuelas, hospitales, tiendas, lotes o grupos experimentales
9. Datos parcialmente observados: datos faltantes, censura en supervivencia y contrafactuales en inferencia causal
10. Introducción aplicada a modelos flexibles: splines, procesos gaussianos, mezclas, redes neuronales bayesianas y modelos no paramétricos bayesianos
11. Taller final con datos reales: formulación de la pregunta, elección de la verosimilitud, *priors*, ajuste, diagnóstico, mejora del modelo y comunicación de resultados
12. Comunicación bayesiana aplicada: interpretar efectos, riesgos, predicciones, limitaciones y sensibilidad de los resultados de forma transparente y reproducible en R

MÓDULO 8: ANÁLISIS DE CURVAS, SEÑALES Y DATOS FUNCIONALES CON R

PRESENTACIÓN Y OBJETIVOS

En muchos problemas reales los datos no aparecen como observaciones aisladas, sino como **curvas, señales, trayectorias o perfiles** registrados a lo largo del tiempo, del espacio, de una longitud de onda o de otra variable continua. Este tipo de datos funcionales aparecen frecuentemente en ciencias de la salud, biología, climatología, química, ingeniería, economía y otros ámbitos aplicados. Algunos ejemplos habituales son electrocardiogramas, curvas de crecimiento, perfiles de actividad, registros de temperatura, espectros de infrarrojos, señales fisiológicas o curvas de consumo.

Este módulo ofrece una **introducción práctica al análisis de datos funcionales** con R. La idea central es representar cada unidad de observación como una función, lo que permite analizar no solo valores puntuales, sino **la forma completa de la curva** o señal. Este enfoque introduce **diferencias importantes** respecto al análisis estadístico clásico y requiere herramientas específicas para representar, suavizar, visualizar, comparar y modelizar este tipo de datos.

A lo largo del curso se presentarán de forma progresiva los conceptos y métodos fundamentales del análisis de datos funcionales, con especial atención a su aplicación en problemas reales. Se abordarán la representación y suavizado de curvas, la visualización de muestras funcionales, las nociones de distancia y profundidad, la detección de patrones y observaciones atípicas, la reducción de dimensión y algunas técnicas de aprendizaje automático para datos funcionales.

Las sesiones combinarán fundamentos estadísticos, ejemplos desarrollados en R y ejercicios guiados, con el objetivo de que los participantes adquieran una visión clara de **cuándo tiene sentido tratar los datos como funcionales, qué ventajas aporta este enfoque y cómo aplicar las principales herramientas disponibles en R** para analizar curvas y señales en sus propios contextos de investigación.

INSCRIPCIÓN Y DATOS PRÁCTICOS

MÓDULO	Análisis de curvas, señales y datos funcionales con R
Docentes	Martín Sánchez y José Luis Torrecilla (Dpto. Matemáticas, UAM) jose Luis.torrecilla@uam.es
Fechas del Curso	10, 11 y 12 de Noviembre de 2026 Horario: 15:00 – 20:00
Inscripciones	Hasta el 6/Nov/2026 en http://matriculas.fuam.es/matriculauam/Convocatorias.action
Lugar de celebración	Universidad Autónoma de Madrid Ctra. Madrid-Colmenar Viejo (M-607), km 15. Madrid. Facultad de Ciencias (C/ Francisco Tomás y Valiente, 7) (por confirmar) http://www.uam.es
Nº Plazas	16 plazas. Mínimo de 8 asistentes para impartirse el curso
Precio	180€ (162€ estudiantes con matrícula en vigor, AlumniUAM+, AlumniUAM+Plus)

	y Amigos de la UAM)
Créditos	1 ECTS
Requisitos previos	No se requiere experiencia previa en análisis de datos funcionales. Para el máximo aprovechamiento del módulo, es conveniente tener conocimientos básicos de R y de modelos estadísticos. Se trabajará con ejemplos de curvas, señales o trayectorias, prestando atención tanto a la implementación en R como a la interpretación de los resultados. En caso de duda contacten con los profesores del módulo.
Observaciones	Los contenidos teóricos, conjuntos de datos y código necesarios para el seguimiento del módulo se harán accesibles en documentos en red. Los participantes deben traer su propio ordenador portátil. Se enviarán indicaciones sobre el software a instalar antes del inicio del módulo.

DESCRIPCIÓN DETALLADA

El módulo se organiza en 3 sesiones de unas 5 horas cada una.

El curso introduce de forma progresiva las principales herramientas para analizar datos observados en forma de curvas, señales o trayectorias. Las sesiones combinarán explicaciones conceptuales, ejemplos desarrollados en R y ejercicios guiados, utilizando conjuntos de datos que permitan ilustrar situaciones habituales en investigación biomédica, experimental y científica.

Se prestará especial atención a la representación de los datos, el suavizado de observaciones discretas, la visualización de muestras funcionales, la comparación entre curvas, la detección de patrones y observaciones atípicas, y la aplicación de técnicas de reducción de dimensión y aprendizaje automático. El objetivo es que los participantes comprendan qué aporta el enfoque funcional frente a un análisis multivariante clásico y sepan aplicar herramientas básicas de análisis funcional de datos con R.

1. Introducción a las curvas, señales y datos funcionales
2. Representación de datos funcionales y suavizado
3. Principales librerías en R: *fda*, *fda.usc*, *tidyfun*
4. Descripción de conjuntos de curvas
5. Visualización, detección de atípicos y reducción de dimensión
6. La noción de cercanía: distancias, métricas, semi-métricas, derivadas...
7. La noción de centralidad y las medidas de profundidad
8. Aprendizaje automático con datos funcionales: clasificación, regresión y clustering

MÓDULO 9: MACHINE LEARNING CON R Y TIDYMODELS

PRESENTACIÓN Y OBJETIVOS

En este módulo se han reunido una serie de técnicas procedentes de distintos campos como el **análisis multivariante, el aprendizaje estadístico, la inteligencia artificial y técnicas de machine learning**. Se han seleccionado las **técnicas predictivas más importantes** y cuyo uso está más extendido: regresión lineal y logística, métodos penalizados (LASSO), métodos basados en árboles (CART), support vector machines (SVM), redes neuronales, random forest.

El módulo incluye los aspectos más importantes para la **construcción de modelos predictivos**: pre-procesamiento y descripción básica de los datos, **optimización de los parámetros** involucrados en cada una de las técnicas, **evaluación de la capacidad predictiva** de los modelos **mediante técnicas de remuestreo (validación cruzada y bootstrapping)**, selección de variables que van a formar parte del modelo, comportamiento de cada técnica en **problemas de alta dimensión**, etc.

Todas las técnicas se verán con los paquetes originales de R, y también se explicará detalladamente cómo se pueden desarrollar con **tidymodels**, el nuevo metapaquete para la estandarización del modelado de datos, con la filosofía de **tidyverse**

INSCRIPCIÓN Y DATOS PRÁCTICOS

MÓDULO	Machine Learning con R y tidymodels
Docentes	Jesús Herranz Valera (Bioestadístico de GEICAM) jherranzvalera@gmail.com
Fechas del Curso	16, 17, 18, 19 y 20 de Noviembre de 2026. Horario: 9:00 – 18:00 los 16, 17, 18 y 19; y de 9:00 – 14:00 el día 20
Inscripciones	Hasta el 12/Nov/2026 en http://matriculas.fuam.es/matriculauam/Convocatorias.action
Lugar de celebración	Universidad Autónoma de Madrid Ctra. Madrid-Colmenar Viejo (M-607), km 15. Madrid. Facultad de Ciencias (C/ Francisco Tomás y Valiente, 7) (por confirmar) http://www.uam.es
Nº Plazas	16 plazas. Mínimo de 8 asistentes para impartirse el curso
Precio	360€ (324€ estudiantes con matrícula en vigor, AlumniUAM+, AlumniUAM+Plus y Amigos de la UAM)
Créditos	3 ECTS
Requisitos previos	Para el máximo aprovechamiento del módulo, son necesarios conocimientos de R y de Estadística, incluidos modelos de regresión (en caso de duda contacten con los profesores del módulo)
Observaciones	Se entregará un documento con todas las diapositivas del curso y todos los scripts y ficheros de datos Es obligatorio que los participantes traigan su propio ordenador portátil , y que se haya instalado el software R previamente

DESCRIPCIÓN DETALLADA

El módulo está dividido en 21 sesiones, con una duración variable de entre 1 y 2 horas, dependiendo de la importancia del tema.

En cada sesión se explican los conceptos estadísticos teóricos ilustrados desde un punto de vista práctico con ejemplos realizados con R, explicando los paquetes y funcionalidades más importantes del software para cada técnica. Finalmente, se proponen ejercicios a los alumnos para que puedan practicar los conocimientos adquiridos.

1. Introducción a Machine Learning. Aprendizaje Supervisado y No Supervisado. Modelos Predictivos. Clasificación y Regresión
2. Introducción a *tidyverse* y *tidymodels*
3. Análisis descriptivo. Pre-procesamiento de datos
4. Regresión Lineal. Métodos de Regresión por pasos
5. Regresión Logística
6. Análisis Discriminante Lineal (LDA). Clasificador Naïve Bayes
7. Construcción de modelos con *tidymodels*
8. Medidas de Evaluación de Modelos. Medidas de capacidad predictiva. Curvas ROC
9. Técnicas de Evaluación de Modelos. Validación cruzada. Muestras Bootstrap
10. Regresión sobre Componentes Principales. Mínimos Cuadrados Parciales (PLS)
11. Métodos de Regresión Penalizados. Ridge Regression. LASSO
12. Métodos basados en Vecindad y Núcleos. k-NN
13. Redes Neuronales (NN)
14. Support Vector Machines (SVM)
15. Árboles de Regresión y Clasificación (CART)
16. Multiclasificadores. Bagging
17. Random Forest (RF)
18. Boosting
19. Caso práctico: Comparación de Modelos Predictivos
20. Caso práctico: Evaluación de la Significación Estadística de un Clasificador mediante Test de Permutaciones